Supplemental Figures

Multi-omics factor analysis reveals a *RAS* mutant-like metabolically distinct AML subgroup resistant to venetoclax

Dominique Sternadt[1], Shanna M. Hogeling[1], Lieve Oudejans[1], Nikita La Rose[1], Gerwin Huls[1], Diego A Pereira-Martins[1], Jan Jacob Schuringa[1#]

[1]Department of Hematology, University Medical Centre Groningen, University of Groningen, Groningen, the Netherlands. #lead contact: j.j.schuringa@umcg.nl

Supplemental Figure Legends

Supplemental Figure 1. Model selection and validation of the robustness of MOFA .

(A) Lines represent the total variance explained (R^2) per view across different factor numbers (K), using MOFA2's variance decomposition. (B) Heatmap depicting the minimum absolute correlation (min|r|) between the factor scores from the full model and those from models retrained after leaving out one (LOO) sample at a time. Higher values indicate factors that are insensitive to removal of any single sample. (C) Absolute Pearson correlations between factor scores from the K=5 (y-axis) and K=6 (x-axis) MOFA models. Correlations were computed across shared samples to assess overlap between the two models. (D) Heatmap showing absolute Pearson correlations between factor scores from the reference K=5 model and a subset of additional runs trained with distinct random seeds. Block-diagonal structure indicates consistent recovery of the same factors across seeds.

Supplemental Figure 2. MOFA factor top loadings and gene set enrichment analysis.

Bubble plots for gene set enrichment analysis (GSEA) summarizing pathway associations for each MOFA factor using the proteomics view and the (A) C2 and (B) C5 gene sets. For every factor, proteins are ranked by their signed MOFA weight (loading). Dot color shows the normalized enrichment score and size encodes -log10(p). Displayed pathways are the top 15 terms that are significantly enriched (FDR < 0.05) in more than one factor, emphasizing biological differences across factors. Heatmaps depicting scaled weights per factor showing all the drugs from the (C)

metabolic and (D) epigenetic drug screen with corresponding culture conditions and time of treatment.

Supplemental Figure 3. Validation of weighted sum of expression (WSE) approach and extrapolation to external cohorts.

(A) Violin plots depicting the distribution of WSE (z-scores) per factor using the proteomics MOFA weights. WSE was applied to the original training matrix in which missing values were imputed using k-nearest neighbors and z-scored per gene. (B) Scatter plots showing the Spearman correlation between WSE scores and the original MOFA factor scores per sample. Correlation (rho) and p-values are depicted in the plots. (C) Stacked bars showing for each external dataset the number of MOFA signature features (n=7992) that were present after gene matching. Percentages above bars indicate the proportion of signature features available for score computation in each cohort. Present and missing features are color coded in the graph. (D) Violin plots show per-factor WSE z-scores in the TCGA[1] (transcriptomics) and Cheng's [2] (proteomics) cohort. (E) Bubble plot summarizing associations between per-factor WSE scores and gene mutation status in Cheng's and TCGA datasets. For each factor x gene, a Wilcoxon rank-sum test compares WSE scores in mut vs wt samples (only genes with ≥3 mutants tested). Dot size reflects $-\log_{10}$(FDR) (Benjamini-Hochberg per factor), and dot color encodes the scaled mean difference. Heatmap of Spearman correlations between factor signature z-scores and drug sensitivity data across inhibitors for (F) Cheng (IC50) and (G) FPMTB[3] (-sDSS) datasets. For each factor, correlations are computed per inhibitor (minimum n ≥ 8), with BH-FDR per factor. Significance stars represent q < 0.05 (*), < 0.01 (**), < 0.001 (***).

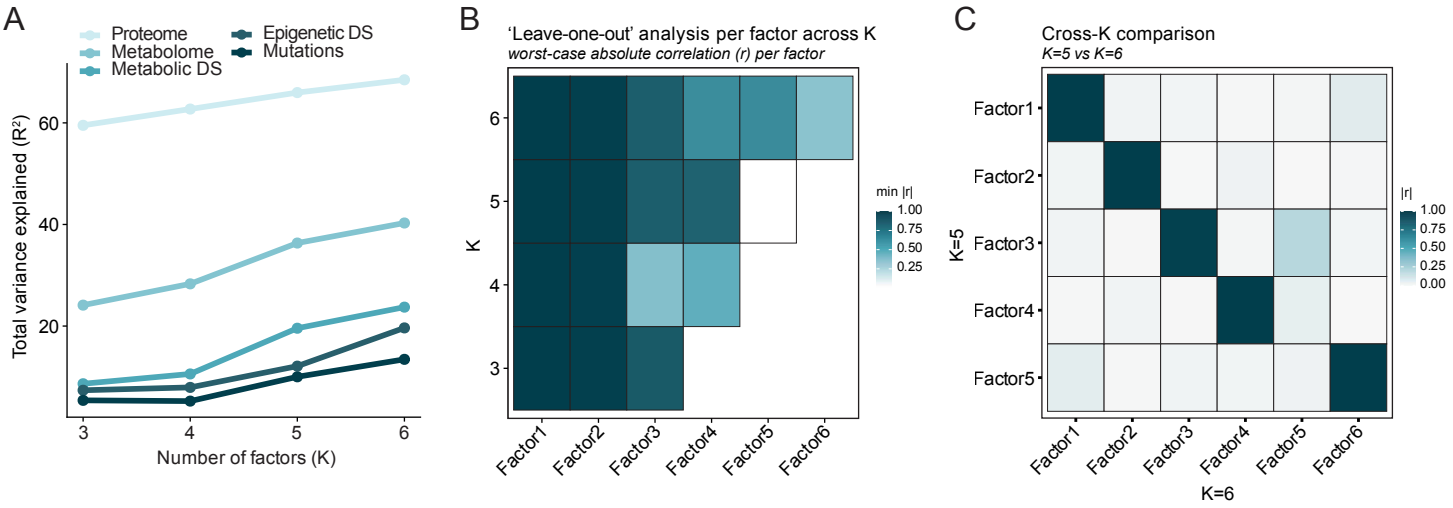
Supplemental Figure 4. Factor 2 associates with differentiation and a *RAS*-mutant-like transcriptional program.

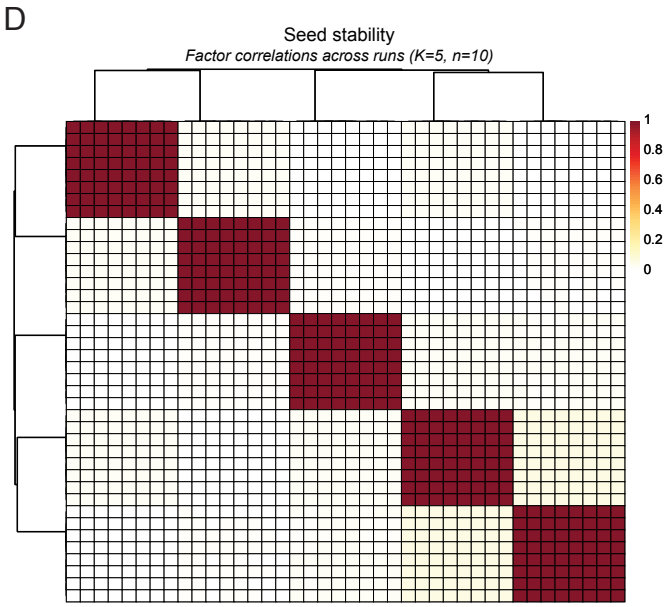
(A) Heatmap of covariate-factor associations from the MOFA model, showing -log10(p-value) for correlations between each latent factor and the indicated clinical/phenotypic covariates. (B) Boxplots of Factor 2 scores across FAB classes in the BeatAML and TCGA cohorts.[1,4] (C) Boxplots showing Factor 2 scores in *PTPN11* wild-type versus mutant samples across four independent AML cohorts, clearly depicted in the graphs. Statisical analysis was done using Wilcoxon test and p-values are stated in the plot. (D) Heatmap depicting single sample gene set enrichment analysis (ssGSEA, z-scored) using the BeatAML cohort for selected terms and samples ordered by Factor 2, with venetoclax response and RAS genotype shown as top annotations. Some of the gene sets names were shortened to simplify visualization. (E) Box plots depicting ssGSEA scores for three gene sets associated with RAS pathways using the FPMTB and TCGA cohorts stratified by *RAS* genotype and Factor 2 high (Factor 2^{high}) vs low (Factor 2^{low}) (Factor 2 > 0 vs ≤ 0). Statistical analysis between groups were assessed using a Kruskal-Wallis test and post hoc pairwise Wilcoxon tests. Adjusted p-values (Benjamini-Hochberg) are depicted in the plots. Volcano plot of per-gene Fisher's exact tests comparing mutation frequency using the BeatAML cohort (excluding *KRAS/NRAS*), showing log2(OR) versus -log10(BH-adjusted q-value) in (F) only RAS^{wt} samples comparing Factor 2^{high} and Factor 2^{low} or (G) only Factor 2^{high} samples (Factor 2 > 0) comparing RAS^{wt} and RAS^{mut}. (H) PCA of transcriptomic profiles colored by *RAS* genotype and Factor 2^{high} and Factor 2^{low} illustrating clustering structure in the BeatAML dataset. (I) Ranked loading plot of the Factor 2 metabolite weights (loadings) from the MOFA model, highlighting metabolites contributing most strongly to the Factor 2 axis.

References

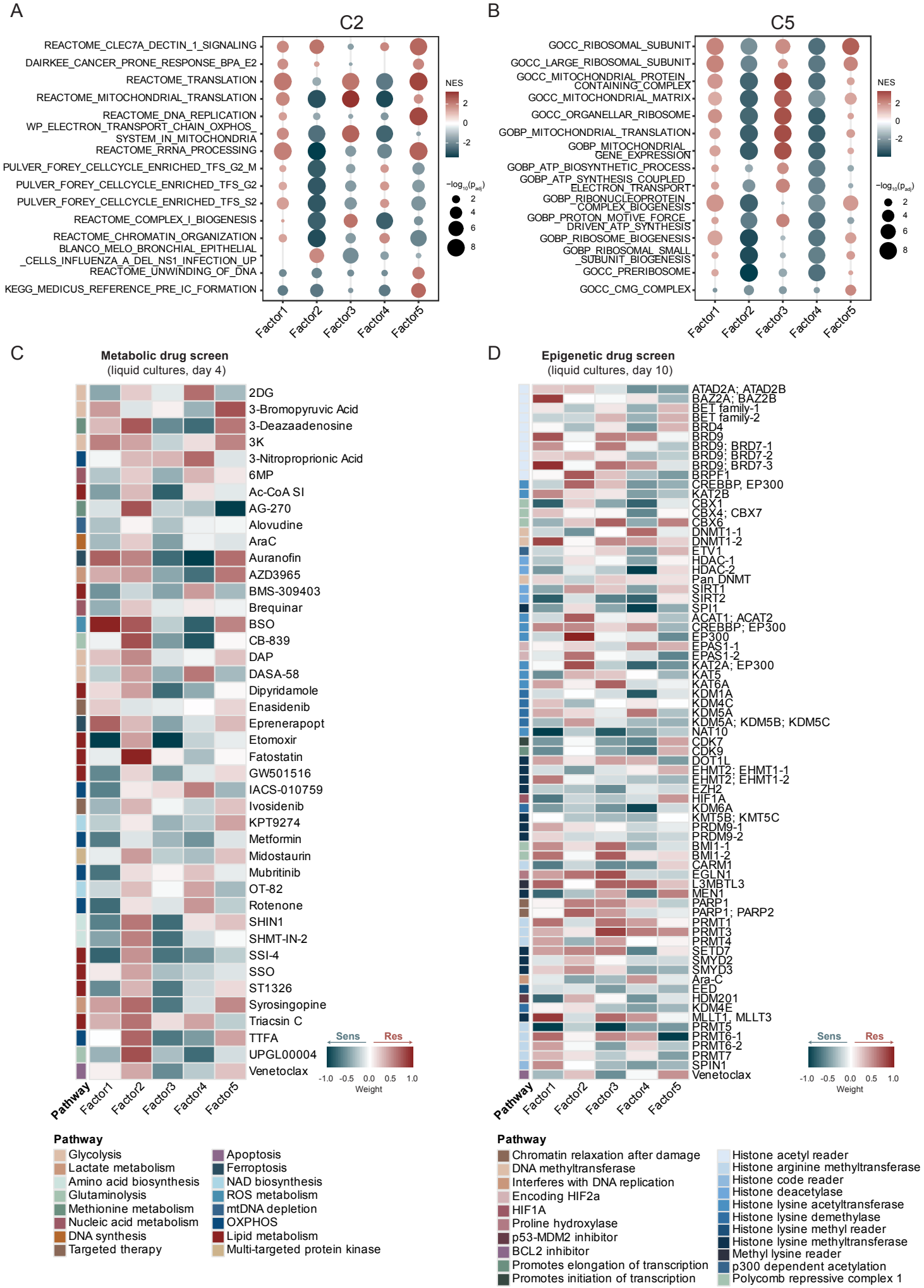
1. Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia. *New England Journal of Medicine* **368**, 2059–2074 (2013).
2. Cheng, W.-Y. *et al.* Proteomic subtypes enrich current acute myeloid leukemia nomenclature and reflect intrinsic pathogenesis alongside aging. *Blood* **146**, 2681–2695 (2025).
3. Malani, D. *et al.* Implementing a Functional Precision Medicine Tumor Board for Acute Myeloid Leukemia. *Cancer Discov.* **12**, 388–401 (2022).
4. Bottomly, D. *et al.* Integrative analysis of drug response and clinical outcome in acute myeloid leukemia. *Cancer Cell* **40**, 850-864.e9 (2022).

Supplemental Figure 1

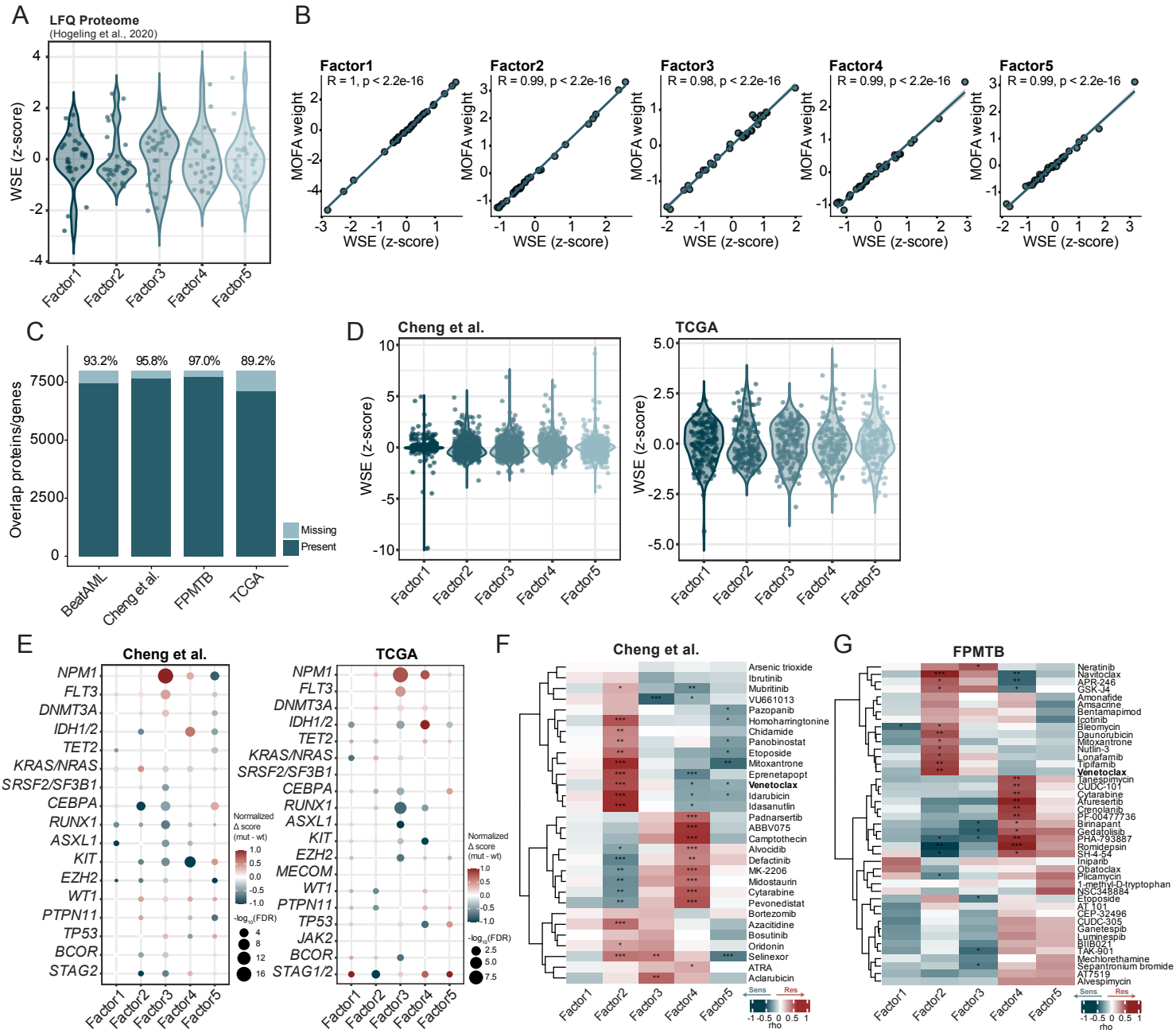




Supplemental Figure 2



Supplemental Figure 3



Supplemental Figure 4